TRANSFORM LEARNING MRI WITH GLOBAL WAVELET REGULARIZATION

A. Korhan Tanc

Department of EEE
Kirklareli University
Kayali, 39100, Kirklareli, Turkey
korhan.tanc@kirklareli.edu.tr

Ender M. Eksioglu

Department of ECE
Istanbul Technical University
Maslak, 34469, Istanbul, Turkey
eksioglue@itu.edu.tr

ABSTRACT

Sparse regularization of the reconstructed image in a transform domain has led to state of the art algorithms for magnetic resonance imaging (MRI) reconstruction. Recently, new methods have been proposed which perform sparse regularization on patches extracted from the image. These patch level regularization methods utilize synthesis dictionaries or analysis transforms learned from the patch sets. In this work we jointly enforce a global wavelet domain sparsity constraint together with a patch level, learned analysis sparsity prior. Simulations indicate that this joint regularization culminates in MRI reconstruction performance exceeding the performance of methods which apply either of these terms alone.

Index Terms— Magnetic resonance, Image reconstruction, Sparsity, Transform learning, Compressed Sensing

1. INTRODUCTION

Using sparsity as a regularizer for ill-conditioned inverse problems has been an active research area in the last decade. Sparse regularization and compressed sensing (CS) ideas have also been applied to image reconstruction in Magnetic Resonance Imaging (MRI). In the pioneering work [1], an equivalent of the following optimization problem has been introduced to regularize the MRI reconstruction problem.

$$\min_{\mathbf{x}} \tfrac{1}{2}\|\boldsymbol{\mathcal{F}}_u\mathbf{x}-\mathbf{y}\|_2^2+\rho_1\|\boldsymbol{\Phi}\mathbf{x}\|_1+\rho_2\|\mathbf{x}\|_{\mathrm{TV}}. \qquad (1)$$

Here, $\mathbf{x}\in\mathbb{C}^N$ is the reconstructed MR image in vectorized form. $\boldsymbol{\mathcal{F}}_u$ is the undersampled Fourier transform operator which realizes the conversion from the vectorized image to the k-space. $\mathbf{y}=\boldsymbol{\mathcal{F}}_u\mathbf{x}^\star+\boldsymbol{\eta}\in\mathbb{C}^\kappa$ is the observation vector in the k-space, where $\mathbf{x}^\star$ is the true underlying image and $\boldsymbol{\eta}$ is the additive noise. The ratio κ/N quantifies the severity of undersampling. $\|\cdot\|_1$ denotes the ℓ_1 norm for the argument vector. $\boldsymbol{\Phi}$ is a sparsifying operator applied to the reconstructed image vector. Throughout this paper we will assume it to be a square wavelet transform. $\|\cdot\|_{\mathrm{TV}}$ is the Total Variation (TV) norm of the argument vector. TV is another common sparse regularizer which has been shown to be beneficial in many image processing applications. In the original Sparse MRI algorithm [1], a nonlinear conjugate gradient method is used for the solution of (1). There are several other approaches for solving (1) or its variants, including the RecPF [2] and FCSA [3] algorithms. In these algorithms operator and variable splitting methods are applied to solve the composite regularization problem (1).

Examplar or patch based methods have also been very popular for sparsity based image processing. Dictionary learning (DL) based synthesis sparsity methods [4] or analysis sparsity [5] based analysis operator learning methods have readily been applied in regularizing corpus of patches extracted from images in a variety of image processing applications. Dictionary learning from patches has been applied to MRI reconstruction in [6] in an algorithm called as the DLMRI. The same authors have developed a novel model for analysis sparsity operator learning, which they have called as sparsifying transform learning (TL) [7]. Sparsifying transform learning can train analysis sparsity operators in a computationally efficient way when compared with other algorithms [8]. The sparsifying transform learning over image patches has been utilized to regularize the MRI reconstruction problem in [9] resulting in the TLMRI algorithm. The results in [6] and [9] indicate that sparse representation of image patches over learned synthesis dictionaries or learned analysis transforms can come up with state of the art performance results for MRI reconstruction.

As discussed above, methods such as Sparse MRI [1], RecPF [2] and FCSA [3] use (1). Hence, they apply global, image-scale regularization on the reconstructed image $\mathbf{x}$. DLMRI [6] and TLMRI [9] algorithms utilize local, patch-scale regularization via learned dictionaries or transforms. However, they lack a global regularizing term. In this work, we aim to bring these two ends together. Recently, a similar idea has been presented in [10], where the DLMRI framework is modified by adding a global TV cost and also by using a different DL algorithm, namely beta process factor analysis (BPFA). The resulting algorithm is called as the CS-MRI-BPFA [10]. In this work on the other hand, we will introduce a general global sparsifying cost into the TLMRI framework

978-0-9928626-3-3/15/$31.00 ©2015 IEEE

and provide the algorithm for solving this modified problem. Working in the transform learning domain we avoid the rather costly dictionary learning step. The resulting algorithm has an efficient reconstruction step realized via proximal splitting, which is different from the least squares step of TLMRI. Our algorithm is also different from the DLMRI and CS-MRI-BPFA algorithms, because these algorithms regularize the image patches using synthesis sparsity over a learned dictionary. The simulation results indicate that the merger of the global [1] and local [9] regularization terms in our algorithm results in reconstruction performance surpassing the usage of either term alone.

2. TRANSFORM LEARNING MRI FORMULATION

TL framework for analysis sparse representation has been introduced in [7]. The square TL approach has been applied to patchwise regularization of MRI image reconstruction in [9]. The cost function as introduced in [9] can be stated as follows.

$$\text{(P0)} \min_{\mathbf{W},\hat{\boldsymbol{\mathcal{X}}},\boldsymbol{\mathcal{A}},\mathbf{x}} \|\mathbf{W}\hat{\boldsymbol{\mathcal{X}}}-\boldsymbol{\mathcal{A}}\|_F^2+\lambda Q(\mathbf{W})+\tau\|\boldsymbol{\mathcal{R}}(\mathbf{x})-\hat{\boldsymbol{\mathcal{X}}}\|_F^2 + \eta\|\boldsymbol{\mathcal{F}}_u\mathbf{x}-\mathbf{y}\|_2^2\,, \quad \text{s.t. } \|\boldsymbol{\alpha}_j\|_0 \le s_j\ \forall j=1\dots M. \quad (2)$$

$\|\cdot\|_F$ is the Frobenius matrix norm, and $\|\cdot\|_0$ denotes the ℓ_0 pseudo-norm. $\mathbf{W} \in \mathbb{C}^{n\times n}$ is the learned square transform which enforces the transform sparsity in the vectorized patch domain. $\hat{\boldsymbol{\mathcal{X}}} \in \mathbb{C}^{n\times M}$, and its columns $\hat{\mathbf{x}}_j \in \mathbb{C}^n$ denote vectorized 2D patches of size $\sqrt{n}\times\sqrt{n}$. The patches stored in $\hat{\boldsymbol{\mathcal{X}}}$ are approximations for the patches of the reconstructed image $\mathbf{x}$, where the $\hat{\boldsymbol{\mathcal{X}}}$ patches also satisfy the transform sparsity model as dictated by the learned transform $\mathbf{W}$. $\boldsymbol{\mathcal{A}} \in \mathbb{C}^{n\times M}$ includes the sparse codes for the patches in $\hat{\boldsymbol{\mathcal{X}}}$, with columns $\boldsymbol{\alpha}_j \in \mathbb{C}^n$. $Q(\cdot)$ represents the required penalization term for the learned $\mathbf{W}$ to avoid degenerate solutions. In the square transform setting considered here $Q(\mathbf{W}) = \|\mathbf{W}\|_F^2 - \log|\det \mathbf{W}|$. We use a modified notation here, and $\boldsymbol{\mathcal{R}}$ is a patch generating operator, such that $\boldsymbol{\mathcal{R}}(\mathbf{x})$ becomes a patch matrix of the same size as $\hat{\boldsymbol{\mathcal{X}}}$. Each column of $\boldsymbol{\mathcal{R}}(\mathbf{x})$ is a patch vector extracted from its proper location by $\mathbf{R}_j\mathbf{x}$, $\forall j = 1,\dots,M$. $\mathbf{R}_j \in \{0,1\}^{n\times N}$ is the patch extracting matrix for an individual patch. The allowed overlap between patches and the patch size determine the number of patches M and the required operator $\boldsymbol{\mathcal{R}}$.

The TLMRI algorithm as described by (P0) applies local patch scale regularization via a learned sparsifying transform. The observation fidelity is enforced on the global image using the $\|\boldsymbol{\mathcal{F}}_u\mathbf{x}-\mathbf{y}\|_2^2$ term as usual. The results in [9] indicate that the patch level regularization has performance surpassing state of the art reconstruction algorithms utilizing sparse regularization via a nonadaptive global operator, such as the wavelet plus TV regularization as in [1]. TLMRI has better performance than the earlier DLMRI algorithm, which enforces sparsity on the patches via a learned dictionary [6]. The complexity of TLMRI is much reduced when compared to the DLMRI, because TLMRI avoids the NP-hard sparse representation substeps. Still the TLMRI only applies regularization at the patch level. It misses a global regularizer as used in many recent algorithms, which have come up with exact and efficient algorithms for the reconstruction [3]. In this work we propose to include an additional global image regularization term in the TLMRI framework. Our modified new cost function together with the global regularizer is as follows.

$$\text{(P1)} \min_{\mathbf{W},\hat{\boldsymbol{\mathcal{X}}},\boldsymbol{\mathcal{A}},\mathbf{x}} \|\mathbf{W}\hat{\boldsymbol{\mathcal{X}}}-\boldsymbol{\mathcal{A}}\|_F^2+\lambda Q(\mathbf{W})+\beta\|\boldsymbol{\mathcal{A}}\|_1 + \tau\|\boldsymbol{\mathcal{R}}(\mathbf{x})-\hat{\boldsymbol{\mathcal{X}}}\|_F^2+\eta\|\boldsymbol{\mathcal{F}}_u\mathbf{x}-\mathbf{y}\|_2^2+\upsilon'\|\boldsymbol{\Phi}\mathbf{x}\|_1. \quad (3)$$

When compared with (P0), in (P1) firstly we have replaced the ℓ_0 norm constraints for the $\boldsymbol{\alpha}_j$ with an ℓ_1 constraint via the $\|\boldsymbol{\mathcal{A}}\|_1$ term, with $\|\boldsymbol{\mathcal{A}}\|_1 = \sum_j \|\boldsymbol{\alpha}_j\|_1$. The other and crucial change is the introduction of the $\|\boldsymbol{\Phi}\mathbf{x}\|_1$ term. Simulations will demonstrate that the addition of this global regularizer enhances performance when compared to the TLMRI which lacks this term. We will denote this modified framework as the Globally regularized TLMRI (G-TLMRI). Next we will try to devise an algorithm for the solution of the G-TLMRI cost (P1) and underline its difference from the TLMRI.

3. THE G-TLMRI ALGORITHM

To solve the minimization problem (P1) (3), we utilize the iterative alternating minimization procedure which has been extensively used in dictionary and analysis operator learning algorithms, and also in TLMRI [9]. As in TLMRI, we firstly separate the algorithm into two steps with and without optimization on $\mathbf{x}$. The first stage should solve the following cost with constant $\mathbf{x}$.

$$\text{(P2)} \min_{\mathbf{W},\hat{\boldsymbol{\mathcal{X}}},\boldsymbol{\mathcal{A}}} \|\mathbf{W}\hat{\boldsymbol{\mathcal{X}}}-\boldsymbol{\mathcal{A}}\|_F^2+\lambda Q(\mathbf{W}) + \beta\|\boldsymbol{\mathcal{A}}\|_1+\tau\|\boldsymbol{\mathcal{R}}(\mathbf{x})-\hat{\boldsymbol{\mathcal{X}}}\|_F^2. \quad (4)$$

We will divide (P2) into two in the following form similar to the TLMRI.

$$\text{(P2.1)} \min_{\mathbf{W},\boldsymbol{\mathcal{A}}} \|\mathbf{W}\hat{\boldsymbol{\mathcal{X}}}-\boldsymbol{\mathcal{A}}\|_F^2+\lambda Q(\mathbf{W})+\beta\|\boldsymbol{\mathcal{A}}\|_1. \quad (5a)$$

$$\text{(P2.2)} \min_{\hat{\boldsymbol{\mathcal{X}}},\boldsymbol{\mathcal{A}}} \|\mathbf{W}\hat{\boldsymbol{\mathcal{X}}}-\boldsymbol{\mathcal{A}}\|_F^2+\beta\|\boldsymbol{\mathcal{A}}\|_1+\tau\|\boldsymbol{\mathcal{R}}(\mathbf{x})-\hat{\boldsymbol{\mathcal{X}}}\|_F^2. \quad (5b)$$

(P2.1) is a transform learning problem, whereas (P2.2) can be considered as patchwise denoising. This division decouples (P2) such that only a fraction of the patches can be used for transform learning, whereas all patches can be used for denoising. (P2.1) can be approximately solved using iterative alternation over two steps [7, 8].

$$\text{(P2.1.1)} \min_{\boldsymbol{\mathcal{A}}} \|\mathbf{W}\hat{\boldsymbol{\mathcal{X}}}-\boldsymbol{\mathcal{A}}\|_F^2+\beta\|\boldsymbol{\mathcal{A}}\|_1. \quad (6a)$$

$$\text{(P2.1.2)} \min_{\mathbf{W}} \|\mathbf{W}\hat{\boldsymbol{\mathcal{X}}}-\boldsymbol{\mathcal{A}}\|_F^2+\lambda Q(\mathbf{W}). \quad (6b)$$

In [9] a similar iteration is devised for transform learning, albeit with the ℓ_0 norm replacing the ℓ_1 norm. Both (P2.1.1) and (P2.1.2) have closed form solutions [9]. (P2.1.1) is solved by simple soft thresholding [11]. The solution for (P2.1.2) involves an SVD over the matrix $\mathbf{L}^{-1}\hat{\boldsymbol{\mathcal{X}}}\boldsymbol{\mathcal{A}}^H$, where $\mathbf{L}$ is the solution to $\hat{\boldsymbol{\mathcal{X}}}\hat{\boldsymbol{\mathcal{X}}}^H + \lambda\mathbf{I} = \mathbf{L}\mathbf{L}^H$. Here, $\mathbf{I}$ is the identity matrix of the appropriate size and $(\cdot)^H$ is the Hermitian transpose.

(P2.2) realizes patch denoising and dealiasing. Two alternating steps for (P2.2) become as follows.

$$\text{(P2.2.1)} \min_{\boldsymbol{\mathcal{A}}} \|\mathbf{W}\hat{\boldsymbol{\mathcal{X}}} - \boldsymbol{\mathcal{A}}\|_F^2 + \beta\|\boldsymbol{\mathcal{A}}\|_1. \quad (7a)$$

$$\text{(P2.2.2)} \min_{\hat{\boldsymbol{\mathcal{X}}}} \|\mathbf{W}\hat{\boldsymbol{\mathcal{X}}} - \boldsymbol{\mathcal{A}}\|_F^2 + \tau\|\boldsymbol{\mathcal{R}}(\mathbf{x}) - \hat{\boldsymbol{\mathcal{X}}}\|_F^2. \quad (7b)$$

(P2.2.1) is the same as (P2.1.1), hence it is solved by soft thresholding. (P2.2.2) has a simple least squares solution for fixed $\boldsymbol{\mathcal{A}}$ given by $(\mathbf{W}^H\mathbf{W}+\tau\mathbf{I})^{-1}(\mathbf{W}^H\boldsymbol{\mathcal{A}}+\tau\boldsymbol{\mathcal{R}}(\mathbf{x}))$. In [9], it is advised to guarantee that the denoising error $\|\boldsymbol{\mathcal{R}}(\mathbf{x}) - \hat{\boldsymbol{\mathcal{X}}}\|_F^2$ goes below a certain threshold C after solving (P2.2.2). This translates into solving (P2.2.1) for a sequence of regularization parameters until the required threshold is reached. We have not realized this approach in our setting. Thus we have stuck with a constant regularization parameter β, as opposed to the adaptive s_j parameter of (P0). Eqns. (5-7) have defined the algorithm for solving (P2).

Now, we are at the position to introduce the second main step for the solution of (P1), namely the reconstruction step.

$$\text{(P3)} \min_{\mathbf{x}} \tfrac{1}{2}\|\boldsymbol{\mathcal{F}}_u\mathbf{x}-\mathbf{y}\|_2^2+\tfrac{\tau}{2\eta}\|\boldsymbol{\mathcal{R}}(\mathbf{x})-\hat{\boldsymbol{\mathcal{X}}}\|_F^2+\tfrac{\upsilon'}{2\eta}\|\boldsymbol{\Phi}\mathbf{x}\|_1. \quad (8)$$

We have normalized the regularization parameters as to leave only two of them intact. Now, let us define the operator $\hat{\boldsymbol{\mathcal{R}}}(\hat{\boldsymbol{\mathcal{X}}}) = \big(\sum_j \mathbf{R}_j^T\hat{\mathbf{x}}_j\big)./\mathbf{w}$. Here "./" is an elementwise division. $\hat{\boldsymbol{\mathcal{R}}}$ is the image generating operator which adds patches together in a suitable manner and normalizes the overlapping pixel values with the proper constants in $\mathbf{w}$ as to generate an averaged vectorized image. After this definition, (P3) can be approximately rewritten as follows.

$$\text{(P3}')\ \min_{\mathbf{x}} \tfrac{1}{2}\big(\|\boldsymbol{\mathcal{F}}_u\mathbf{x}-\mathbf{y}\|_2^2+\tau'\|\mathbf{x}-\hat{\boldsymbol{\mathcal{R}}}(\hat{\boldsymbol{\mathcal{X}}})\|_2^2\big)+\upsilon\|\boldsymbol{\Phi}\mathbf{x}\|_1. \quad (9)$$

Now we define two functions $g(\mathbf{x}) = \frac{1}{2}\big(\|\boldsymbol{\mathcal{F}}_u\mathbf{x} - \mathbf{y}\|_2^2 + \tau'\|\mathbf{x} - \hat{\boldsymbol{\mathcal{R}}}(\hat{\boldsymbol{\mathcal{X}}})\|_2^2\big)$ and $f(\mathbf{x}) = \upsilon\|\boldsymbol{\Phi}\mathbf{x}\|_1$. Hence, the cost in (P3′) is of the form $f(\mathbf{x}) + g(\mathbf{x})$, where $g(\mathbf{x})$ is a smooth and differentiable convex function and $f(\mathbf{x})$ is a nonsmooth and nondifferentiable convex function. This problem can be solved very efficiently by proximal splitting methods. We have in particular used the well-known forward-backward splitting algorithm [12]. The forward-backward splitting iterations consist of the below given two steps.

$$\text{(P3.1)} \quad \mathbf{z} = \mathbf{x} - \gamma\nabla g(\mathbf{x}). \quad (10a)$$

$$\text{(P3.2)} \quad \mathbf{x} = \mathbf{x} + \mu(\text{prox}_{\gamma f}(\mathbf{z}) - \mathbf{x}). \quad (10b)$$

Algorithm 1 G-TLMRI Algorithm

Input: Observation, $\mathbf{y} = \boldsymbol{\mathcal{F}}_u\mathbf{x}^\star + \boldsymbol{\eta}$; parameters $\lambda, \beta, \tau, \tau', \upsilon, \gamma, \mu$.

Goal: $\min_{\mathbf{W},\hat{\boldsymbol{\mathcal{X}}},\boldsymbol{\mathcal{A}},\mathbf{x}} \|\mathbf{W}\hat{\boldsymbol{\mathcal{X}}} - \boldsymbol{\mathcal{A}}\|_F^2 + \lambda Q(\mathbf{W}) + \beta\|\boldsymbol{\mathcal{A}}\|_1 + \tau\|\boldsymbol{\mathcal{R}}(\mathbf{x})-\hat{\boldsymbol{\mathcal{X}}}\|_F^2+\eta\|\boldsymbol{\mathcal{F}}_u\mathbf{x}-\mathbf{y}\|_2^2+\upsilon'\|\boldsymbol{\Phi}\mathbf{x}\|_1$

Initialize $\mathbf{x} = \boldsymbol{\mathcal{F}}_u^H\mathbf{y}$.
for $i := 1, 2, \ldots$ **do** ▷ main iteration
  Initialize $\hat{\boldsymbol{\mathcal{X}}} = \boldsymbol{\mathcal{R}}(\mathbf{x})$. ▷ denoising starts
  Iterate (6), N_1 times.
  Iterate (7), N_2 times.
  Initialize $\mathbf{x} = \hat{\boldsymbol{\mathcal{R}}}(\hat{\boldsymbol{\mathcal{X}}})$. ▷ reconstruction starts
  Iterate (10), N_3 times.
end for ▷ end of main iteration
Output reconstructed MR image $\mathbf{x}$.

Here, we have utilized constant values for the step-size γ and relaxation parameter μ. Using the above given definition of the function $g(\cdot)$, its gradient is calculated as $\nabla g(\mathbf{x}) = \boldsymbol{\mathcal{F}}_u^H(\boldsymbol{\mathcal{F}}_u\mathbf{x} - \mathbf{y}) + \tau'(\mathbf{x} - \hat{\boldsymbol{\mathcal{R}}}(\hat{\boldsymbol{\mathcal{X}}}))$. $\boldsymbol{\mathcal{F}}_u^H$ is the adjoint operator of $\boldsymbol{\mathcal{F}}_u$, and it realizes zero-filled reconstruction. The $\text{prox}_h(\cdot)$ operator denotes the proximal operator for the function h. For $f(\mathbf{x}) = \upsilon\|\boldsymbol{\Phi}\mathbf{x}\|_1$, $\text{prox}_{\gamma f}(\cdot)$ is realized by soft thresholding in the transform ($\boldsymbol{\Phi}$) domain with the proper constant and consequently taking an inverse transform. In our simulations we have utilized only a few (e.g. 5) iterations of the forward-backward algorithm each time it is invoked, and we have initialized it with $\mathbf{x} = \hat{\boldsymbol{\mathcal{R}}}(\hat{\boldsymbol{\mathcal{X}}})$. We should note that the modified reconstruction step (P3′) and its solution constitute the main difference of our proposed method from the TLMRI. The reconstruction step of the TLMRI which lacks the global regularization term is simply solved by least squares.

With the reconstruction step we have finalized the G-TLMRI algorithm. Eqns. (3-10) and the corresponding solutions constitute the G-TLMRI algorithm. We give a condensed description of the G-TLMRI algorithm in Alg.1.

4. SIMULATION RESULTS

In this section we compare the reconstruction performance of the new G-TLMRI algorithm with the original TLMRI [9], the DLMRI[1] [6] and the FCSA[2] [3] algorithms. The transform learning and denoising parts of the G-TLMRI and TLMRI are realized with the exact same parameters and methodology using (P2). This facilitates a fair comparison and underlines the contribution of the novel global regularization term in G-TLMRI. The patch size in all three patch based algorithms is 6×6. We employ maximally overlapping patches for denoising, but only a fraction (200×36) of these

[1] http://www.ifp.illinois.edu/~yoram/DLMRI-Lab/DLMRI.html
[2] http://ranger.uta.edu/~huang/R_CSMRI.htm

patches are used for training. The inner iteration numbers are $N_1 = N_2 = 10$ and $N_3 = 5$. The parameters for each method are approximately optimized for the best SNR result. For TLMRI and G-TLMRI $\lambda = 10^5$, $\beta = 0.02$ and $\tau = 1$. For G-TLMRI $\tau' = 10^{-3}$, $\upsilon = 10^{-4}, \gamma = 10, \mu = 0.1$. For FCSA which solves (1), $\rho_1 = 10^{-1}$ and $\rho_2 = 5 \times 10^{-3}$. DLMRI is realized using its publicly available code with the parameters as optimized in [6]. The simulations are repeated with the same parameters for two separate MR images of size (256×256). The downsampling ratio for $\boldsymbol{\mathcal{F}}_u$ is $\kappa/256^2 = 0.25$ (4 fold downsampling) with a random sampling mask in a noiseless setting. The simulations were realized in Matlab on a computer with an Intel i7 CPU at 2.4GHz, 12GB memory and 64-bit operating system. The signal-to-noise ratio (SNR) is calculated as follows: SNR $= 10\log_{10}\big(\text{var}(\mathbf{x}^\star)/\text{MSE}(\mathbf{x})\big)$, where var$(\cdot)$ and MSE$(\cdot)$ denote the variance and mean square error, respectively.

Fig.1 shows a sample undersampling mask in the k-space, and also the original MR images. Figs. 2 and 3 detail the reconstructed images and SNR curves for the brain and shoulder MR images, respectively. The local patch regularization based TLMRI has better performance than the global regularization based FCSA and also the DLMRI for both images. However, it is also evident that the global regularization term introduced in the novel G-TLMRI algorithm has a clear effect on performance, as the G-TLMRI has almost a 2dB gain over the TLMRI. Hence, we can state that the use of both the local patch regularization and global sparse regularization terms together has resulted in performance gain when compared to the use of these terms alone. The execution times per iteration for the different methods are given in Table 1. The non-patch, globally regularizing FCSA is clearly the fastest. G-TLMRI is only very slightly slower than the TLMRI. DLMRI is the slowest due to its use of NP-hard sparse representation steps.

Table 1: Time per iteration for the algorithms.

Algorithm	G-TLMRI	TLMRI	DLMRI	FCSA
Time (sec)	1.69	1.55	39.4	0.025

5. CONCLUSIONS

We have presented a new algorithm called as G-TLMRI for MRI reconstruction. G-TLMRI algorithm builds upon the patch level sparsification approach of the previously introduced TLMRI. G-TLMRI introduces a sparsity based but this time global regularizer into the TLMRI framework. The performance of G-TLMRI is compared with TLMRI and two other recent methods from the literature. Simulation results indicate that G-TLMRI has reconstruction performance exceeding all three of these competing methods. Hence, we can state that the combination of the local and global regularization terms as in G-TLMRI has some merit which deserves further study.

REFERENCES

[1] M. Lustig, D. Donoho, and J.M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magnetic Resonance in Medicine*, vol. 58, no. 6, pp. 1182–1195, 2007.

[2] J. Yang, Y. Zhang, and W. Yin, "A fast alternating direction method for TVL1-L2 signal reconstruction from partial Fourier data," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 2, pp. 288–297, April 2010.

[3] J.Huang, S.Zhang, and D.Metaxas, "Efficient MR image reconstruction for compressed MR imaging," *Medical Image Analysis*, vol. 15, no. 5, pp. 670 – 679, 2011.

[4] E.M. Eksioglu, "Online dictionary learning algorithm with periodic updates and its application to image denoising," *Expert Systems with Applications*, vol. 41, no. 8, pp. 3682 – 3690, 2014.

[5] S. Nam, M.E. Davies, M. Elad, and R. Gribonval, "The cosparse analysis model and algorithms," *Applied and Computational Harmonic Analysis*, vol. 34, no. 1, pp. 30 – 56, 2013.

[6] S. Ravishankar and Y. Bresler, "MR image reconstruction from highly undersampled k-space data by dictionary learning," *IEEE Trans. Med. Imag.*, vol. 30, no. 5, pp. 1028–1041, May 2011.

[7] S. Ravishankar and Y. Bresler, "Learning sparsifying transforms," *IEEE Trans. Signal Process.*, vol. 61, no. 5, pp. 1072–1086, 2013.

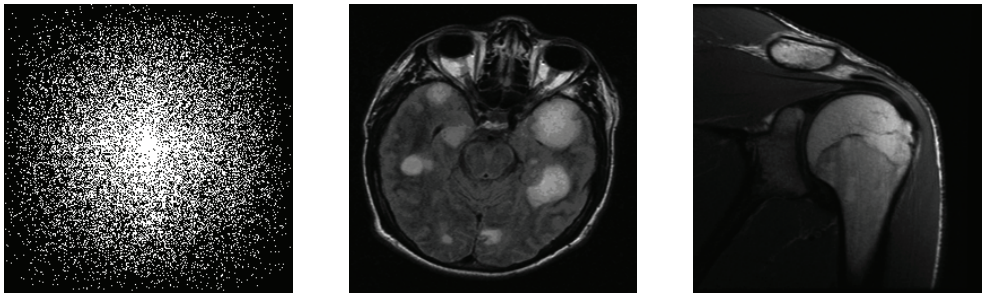
[8] E.M. Eksioglu and O. Bayir, "K-SVD meets transform learning: Transform K-SVD," *IEEE Signal Process. Lett.*, vol. 21, no. 3, pp. 347–351, March 2014.

[9] S. Ravishankar and Y. Bresler, "Sparsifying transform learning for compressed sensing MRI," in *2013 IEEE 10th International Symposium on Biomedical Imaging (ISBI)*, April 2013, pp. 17–20.

[10] Y. Huang, J. Paisley, Q. Lin, X. Ding, X. Fu, and X.P. Zhang, "Bayesian nonparametric dictionary learning for compressed sensing MRI," *IEEE Trans. Image Process.*, vol. 23, no. 12, pp. 5007–5019, Dec 2014.

[11] E.M. Eksioglu and O. Bayir, "Overcomplete sparsifying transform learning algorithm using a constrained least squares approach," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, May 2014, pp. 7158–7162.

[12] P.L. Combettes and J.C. Pesquet, "Proximal splitting methods in signal processing," in *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, Optimization and Its Applications, pp. 185–212. Springer New York, 2011.



a) b) c)

Fig. 1: Sampling mask in k-space with 4-fold undersampling (a), the original MRI test images (b,c).

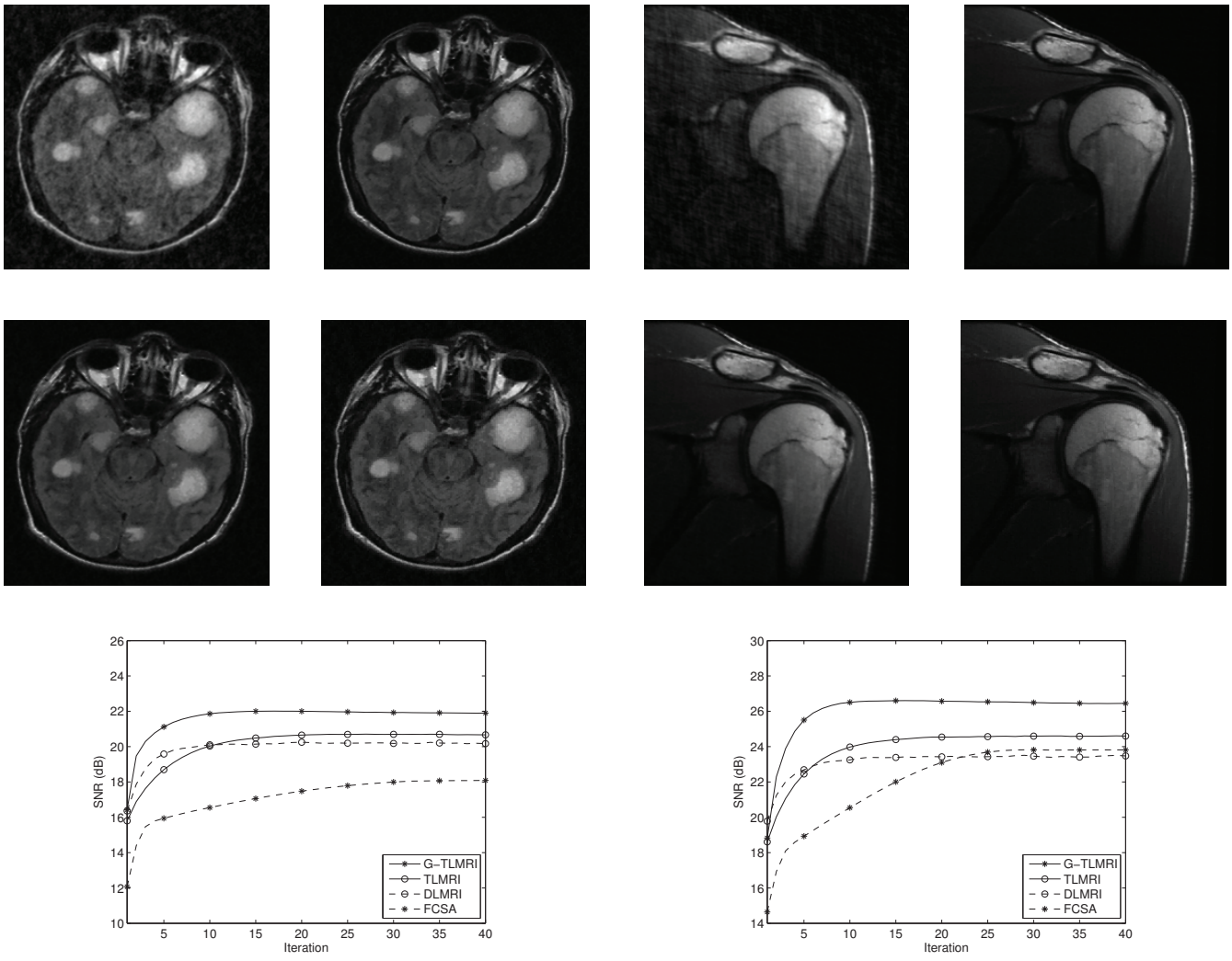


Fig. 2: Brain image results. First row: Zero-filling reconstruction (left), G-TLMRI reconstruction (right). Second row: TLMRI reconstruction (left), FCSA reconstruction (right). Third row: SNR versus iteration.

Fig. 3: Shoulder image results. First row: Zero-filling reconstruction (left), G-TLMRI reconstruction (right). Second row: TLMRI reconstruction (left), FCSA reconstruction (right). Third row: SNR versus iteration.